MSDS 413, Summer 2019, Assignment 4
Stationary Univariate ARMA Models (TS4)

Introduction

Consider the monthly series of Consumer Sentiment of the University of Michigan. This survey series is widely used to indicate the consumer confidence about the U.S. economy. The data are available from FRED of the Federal Reserve Bank of St. Louis and also in the file m-umcsent.txt. The sample period is from January 1978 to August 2013.

The following list defines the Consumer Sentiment variables:

- year: is the year the data were collected
- mon: is the month the data were collected
- day: is the day the data were collected
- VALUE: is the measure of consumer confidence in the U.S. economy

Average (Mean) Duration of Unemployment of the U.S. Department of Labor: Bureau of Labor Statistics monitors the Employment Situation and it is Seasonally Adjusted. The series is monthly values beginning 1948.01.01 to 2014.03.01.

The following list defines the Unemployment Situation variables:

- year: is the year the data were collected
- mon: is the month the data were collected
- day: is the day the data were collected
- Value: is the duration of unemplyoment

Consider the weekly crude oil prices: West Texas Intermediate (WTI), Cushing, Oklahoma. The data are also available from FRED and in the file w-coilwtico.txt. The sample period is from January 3, 1986 to April 2, 2014.

The following list defines the weekly crude oil prices variables:

- Year: is the year the data were collected
- Mon: is the month the data were collected
- Day: is the day the data were collected
- Value: is the measure of consumer confidence in the U.S. economy

Your objective is to explore the time series behavior of these data sets including EDA, modeling, model diagnostics, and interpretation.

Procedure

The following steps are necessary to complete this assignment. Address each and every part and ensure that you cover all the details specified in the questions.

1. **EDA: Consumer Sentiment** (2 points)

   1.1. Conduct a complete EDA on the monthly series of Consumer Sentiment of the University of Michigan.

   1.2. Are there unit roots and if yes, what do they mean?

   1.3. For the first-differenced series, denote the change series by r_t and let $E(r_t) = \mu$. Perform a complete EDA.

   1.4. Compare the differenced data EDA to the undifferenced data EDA.

2. **Duration analysis** (2 points) Unemployment rate is an important macroeconomic series. Equivalent in importance is the duration of unemployment. Consider the mean duration of unemployment is the U.S. from January 1948 to March 2014. The duration is measured in weeks.

   2.1. For the first-differenced data, r_t, test $H_0 : \mu = 0$ versus the alternative $H_a : \mu \neq 0$. What does this mean?

   2.2. Build an AR model for r_t series. Perform model checking using gof = 24. Is the model adequate? Why?

   2.3. Write the equation of your fitted AR model.

   2.4. Fit a seasonal model for the r_t series using the command
   ms <- arima(rt, order=c(2,0,1), seasonal=list(order=c(1,0,1), period=12), include.mean=F)
   Perform model checking using gof = 24. Is the seasonal model adequate? Why?

   2.5. Based on in-sample fitting, which model is preferred? Why?

   2.6. Consider out-of-sample predictions. Use $t = 750$ as the starting forecast origin. Which model is preferred based on the out-of-sample predictions?

3. **Commodity prices modeling** (2 points)

   3.1. Let r_t be the growth rate series, i.e. the first difference of log oil prices. Is there serial correlation in the r_t series? You may use Q(10) to draw the conclusion.

   3.2. Build an AR model for r_t. Check the adequacy of the model, and write the equation of the model to be estimated.

   3.3. Fit another model to r_t using the following command
   m5 <- arima(rt, order=c(3,0,2), include.mean=F)
   This is an ARIMA(3,0,2) model. Write the model equation to be fitted.

   3.4. Based on in-sample fit, which model is preferred?

4. **Report** (1.5 points) For the Commodity prices data, describe to a client or employer what you reccommend. The report requires information from which decisions can be made or actions taken.

Deliverables

See Section Submission Directions below. The assignment deliverables, each in pdf format, are as follows:

- *Only if requested by instructor*
  - The program or script
  - Logs
  - Outputs
- Mandatory
  Data analysis write-up: no programs, logs, or just code outputs.

The data analysis must follow and use the item numbering of each assignment, i.e., use the numbers, say, 1 - 5, with the sub-lettering if used. These deliverables are provided according to the instructions in the Submission Directions section below.

Submission Directions

Title Page

Include a title page with your name and the assignment designation. Leave room for instructor comments.

File Names

The assignment write-up file shall be submitted to Canvas according to the schedule in the syllabus using the item (1) naming convention below. The naming convention is case sensitive. Use letters and numbers as given. **The file name parts have no spaces or other separator charancters**.
TS4Lastname.pdf (submit via Canvas)

The parts are the assignment code, TS4; your lastname with only the first letter capitalized; a period, and lastly, the extension "pdf". Generically,

TS4Lastname.pdf

For example: Suppose your name is Student McStats. Your filename then is:

TS4Mcstats.pdf

The analysis write-up file must be submitted for grading. Each write-up requires a title page for instructor comments. The analysis may use either R or any other statistics package you wish, or if you use more than one package, you must use the germane tables, plots, etc., in a single report. If you use more than one package, differences and similarities should be indicated.

email: jamie.riggs@northwestern.edu

Email *ONLY IF REQUESTED* the program (script), log and output as separate pdf files. The R log and output may be combined. The file names shall be as follows:

- The program or script file names
  - TS4LastnameRprog.pdf
- The log file names
  - TS4LastnameRlog.pdf